

04 D.O



OIPE

2

RAW SEQUENCE LISTING DATE: 01/17/2002  
PATENT APPLICATION: US/10/029,120 TIME: 17:49:50

Input Set : N:\CrF3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

1 <110> APPLICANT: Swanson, Ronald V.  
2 Feldman, Robert A.  
3 Schleper, Christa  
4 <120> TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
5 <130> FILE REFERENCE: DCORP.002A  
6 <140> CURRENT APPLICATION NUMBER: 10/029,120  
7 <141> CURRENT FILING DATE: 2001-12-21  
8 <150> PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
W--> 9 <151> PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
10 <160> NUMBER OF SEQ ID NOS: 123  
11 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
12 <210> SEQ ID NO: 1  
13 <211> LENGTH: 32998  
14 <212> TYPE: DNA  
15 <213> ORGANISM: Cenarchaeum symbiosum  
16 <220> FEATURE:  
17 <221> NAME/KEY: CDS  
18 <222> LOCATION: (7604)...(8908)  
19 <221> NAME/KEY: CDS  
20 <222> LOCATION: (8961)...(9767)  
21 <221> NAME/KEY: CDS  
22 <222> LOCATION: (10545)...(10922)  
23 <221> NAME/KEY: CDS  
24 <222> LOCATION: (13944)...(14612)  
25 <221> NAME/KEY: CDS  
26 <222> LOCATION: (18638)...(20149)  
27 <221> NAME/KEY: CDS  
28 <222> LOCATION: (20554)...(20955)  
29 <221> NAME/KEY: CDS  
30 <222> LOCATION: (20956)...(21834)  
31 <221> NAME/KEY: CDS  
32 <222> LOCATION: (25151)...(26377)  
33 <221> NAME/KEY: CDS  
34 <222> LOCATION: (27535)...(28002)  
35 <221> NAME/KEY: CDS  
36 <222> LOCATION: (28065)...(29483)  
37 <400> SEQUENCE: 1  
38 gatccttgac ctctgcgctt attgcagcca tggactgacc ggccgtcgcc ggctaaataa 60  
39 agtcgaggcc cgcgcctgcgc gctctgccta gccctgtatt atacagtaact cgcactcgca 120  
40 ggtttcgctt tcgtcatgg ttggccatgtc ctggccgtgc gttccggact ggcacttgac 180  
41 gcagatctgc tctatgttgt ctatcgagggt gtacttttg gactgttcaa agccgaagcc 240  
42 gccgttcattt ggttccaaacctt tgaaggcggtc agggccgttta tgccatattt atctttgcg 300  
43 tcggccgggg ggccttcgttgc ccggccgggg ccccccttgg qccggcttccc cggctcttgg 360  
44

ENTERED

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002  
TIME: 17:49:50

Input Set : N:\CrF3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002

TIME: 17:49:50

Input Set : N:\Crf3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

94	ttagatctt	tccaaagggg	tgccggccca	gcccgaactgc	ccacctgcac	atgtccccgg	3360
95	tcttcacgg	gtatggca	ctgcggaaa	tgtctgtgt	tatcatcgcc	tcccccgtac	3420
96	tcccaaaggaa	ccgcggaaa	tgactcccg	atacgtcat	cactccctgc	tatcacaaaa	3480
97	gacaacgtg	cgtataaaact	caacggggtc	tttcttcggcc	gatggaaat	gatggactgt	3540
98	tgtccacat	tatgtggct	cacgggttgc	tagggggaa	cgtggggct	ctggttgttc	3600
99	cattatcgca	ctgcggaaat	taccggacaa	ggatggttgc	tacctaataa	gagtccagat	3660
100	tatccccggc	atccaaagggt	ccttgcgtgc	tttgcggaaa	aggttatgtat	acccggccac	3720
101	ggccaggatcc	aggcactata	catacccttc	cgggatcaga	tgccgtctgt	ttttttttat	3780
102	acatgtggaa	ccccctgttc	actgcaccc	gttgcggca	ttctctatcg	cagctgcagg	3840
103	catccctat	aaactaaggta	caggactata	tttgcggaaa	ccctccgtat	acggatatacc	3900
104	cgtatgcacat	tgatgttacta	aaccaggcga	cctgttgcgg	atatggggta	gaaatgttgcg	3960
105	tttgtatggc	gcaatggttt	ttatgttgc	tttgcggaaa	tttttttttt	ctaaacggaaa	4020
106	ggccactcccg	cttcggggct	gttttcgtca	tttgcggaaa	tttttttttt	ctaaacggaaa	4080
107	gacacgacgaa	cggtatgttgc	ccccctatcc	ggaaaggaaac	ctatcggttt	aaacggctcc	4140
108	tgcggatgttgc	cagaatattat	actgttttttc	catttcgtat	acttcgttgc	ggcagtttcc	4200
109	aggatcgact	atccggggc	tgacggcga	tttgcggaaa	accctttgc	ttatgggttgc	4260
110	ggcgatgttc	acccggatct	gtcttttttc	ccacccggat	cttgcataata	aaatccgttca	4320
111	caggacgtca	ccggcccttc	ttgttccaaat	cactacggcc	acccatccac	gttgcggatcat	4380
112	cacgggttgc	ttgtggatgttgc	ttgttttttc	tttttttttt	tttttttttt	tttttttttt	4440
113	tcgctggca	gttaatgtgt	tacacacttt	tttgcggaaa	tttttttttt	tttttttttt	4500
114	ctctgtgttgc	tttgcggatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4560
115	taatccatgt	cttgggtttaa	cccccttcgg	tttgcggaaa	tttttttttt	tttttttttt	4620
116	catatgttttgc	ttgtgtgttgc	tttgcggaaa	tttttttttt	tttttttttt	tttttttttt	4680
117	cgccgcacccat	tatcatgttgc	tttgcggaaa	tttttttttt	tttttttttt	tttttttttt	4740
118	cgttccgtttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4800
119	cacacaaacat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4860
120	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4920
121	tcacacgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4980
122	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5040
123	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5100
124	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5160
125	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5220
126	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5280
127	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5340
128	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5400
129	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5460
130	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5520
131	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5580
132	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5640
133	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5700
134	ttatcgatgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5760
135	tttgcacccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5820
136	ccggatgttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5880
137	gataacggca	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	5940
138	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	6000
139	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	6060
140	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	6120
141	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	6180
142	ccggatatttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	6240

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002  
TIME: 17:49:50

Input Set : N:\CrF3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

143	tcttttacacttaccaagg acatcttcacg gcacgagctg gcgacggcca tgcaccacct	6300
144	ctcagttgtt ctggtagagt cttcacgttg accttcacac tgctgtctc cggggtaaga	6360
145	ttttctgggt tgactccaat tgaaccccgag gtcttacccc ttgttggtct cccccggccaa	6420
146	ttcttttaag tatcatactt ggtgtactac ttcccaggcg gcaacttaa cggcttccct	6480
147	ggggcactgc actgtcttc aacgcacatg atactggat ttgoattgt tacagctgg	6540
148	actacccggg tatctaatac ggtttgtcc cccacgttcc atccctaccc gtcggacgtg	6600
149	ttcttagata ccgccttcgc cacaggggtt catcgataga tcagaggatt ttaccccttc	6660
150	cttccggatc cctgttactt ctccacactt ctacggctgc agtatttcg gtagccatcg	6720
151	cgttggcgc atatgtttaa ccggaaactt acacggcagg ctacggatgc tttaggccta	6780
152	ataatccctc tgaccaactt gggtgtctgtt ttacccggc cgggtgcacac cagaacttgc	6840
153	ccacccccc ttccgggtt gtttaagac cggtaaaaaga ttcttttgc agaaaaact	6900
154	cggttaacc ttgtgtgtc ttccggcatt gcaaaatttt ctcgcgtctc ggcggccata	6960
155	ggggctgggtt cgtgtcttc gtaaccatctt cggggctctt ctcttcagag cccgtatctg	7020
156	tttagctt ggtggccat taccatcca acaactgtat agacccgacg cccatcttac	7080
157	ggcgataaaat cattttggcc acaaaccatt ccaggatcg tggctatcg gatatttttc	7140
158	tcgtttttttt gaggttatcc cctgttccatg tttagatgg tctatgtt ctcggcgtc	7200
159	tgccttgtat tgcataatcg atctcgatgg tttagatca atccgtatgc agtcgggtcc	7260
160	ggcaggatca accggatcca taatggatt atttttttt ttgttaatgtc gtttgtactt	7320
161	ttggaaatca acagaatgtc cataatcttc acatctcaga tatggaccctt cgatcatacc	7380
162	ctcatcttgtt tgctgtactt ggaggccacg gaaatcaat atgttacat attggatcatt	7440
163	catcgaaacg ggcgttctt ctgtcaatcgat gatcgatgg gcccgttcat gggcatataa	7500
164	accatcgccg atttcgccecc cggcggcagee cgatcagggg cggatctgc ctgtatgtat	7560
165	ggatccggcc ctgtataaaat tatggggggg gggccgtgt gccgttggatc tggaaacgca	7620
166	gtacaggggc aacggccggc gtcgtccggc gatcttgcgg aggttogaataa agtaccacgt	7680
167	cggcggggc acgaccaaca taagggttata cggatcgatg ctgttgcg caaggctcgc	7740
168	gaggccggaa cactctgtcg acgttgacgg gaaacaaatgt gtagactact ggtatggggc	7800
169	ctggagcttc atactggggc acggccggcc gcaatcgatc tggcgatgtt agggccatgt	7860
170	tcgcggccgc tgatgtttttt ccggccgttccatc cggacgacg agtataatctt cggatataa	7920
171	acgcggccgcg gtaacgtgtt cagaaaaagac aaggatcgatc acgttgcggg cggaggccgt	7980
172	catgtatgc gcaacgtgtt ccgcggccgcg tccggccatc taaggccaga aaaaatataag caaaggccga	8040
173	cggcggctgtt caccgggtatc ctgtccgggtt ctgtcaatcg gtcactatgc ctgtatgtt	8100
174	gcccggaaatc gggggggatc tccggacaaatcg gactatataa tccatccgtt acaaacgtat	8160
175	tggaaatgttcc tggatgttc ttggccggc acggacgacg ttggatgtcg tgataatcg	8220
176	gcgcgtctgg ggccggccggc gctgtatacc gggggatgtgg gactatctgc gggccatata	8280
177	ggatgttgtt catcaatgggg ggcggctgtt tgcttcggatc gagatgttgaa cgggttccg	8340
178	gtttaggtttt ggtctggcgatc atgtcgatggc aggttgcggatc cccatataatg tggctctgg	8400
179	caagatgtt gggggggat tccccatagg ggtatgtatgc gcaaaaggccg aggtgtatggaa	8460
180	aatctccaaatc actatatcgatc tgatcaatggc cggacggccg tacatccggc gggccatatt	8520
181	ctctgcacaaatcc cccggccacaa tgacacgggg cggccggccg ctccggggacg tcaaaaaagag	8580
182	aaaggccaca atatacccgaa ggataaactt catggggggc gacgcacagg acaacttc	8640
183	aaagatattt gggaaatgggg tatecgatgttccatc cggaaaggccc tcgtgttca tgactatctt	8700
184	tgttcaatgtt ccgcggccgcg ggttcaattdt tgctgtcgat gggccatgttgc gcatgttga	8760
185	gctgtgtccatc aggttacccatc tggacatgtt caccggggac ggcataatctt ttctggccgg	8820
186	caacgttggg gccatatacg gccggacactt aacggccgc acatccaaatcg tggatccgc	8880
187	atcacagacgcg ttggcggatgg gcttatgttgcg tatacgatccg gaggaaactt tgattatcg	8940
188	ggctgtgtcc cccggggccc atgtactactt tggccatggag cggacccggcc gagctgtgtc	9000
189	gcacggccgcg cttccgttgcg acgaaacaaatcg agttccggc ggcataatgc ctgtacggc	9060
190	aaatctccaaatcc gggacggccgc cagaacacggg gctgtcttc gaaaaaaggggg ctggcccgaa	9120
191	acaggccaaa aaatgtatctt gatcgatca cgtgttttgcg cggctgtctc gagctgtaca	9180

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002  
TIME: 17:49:50

Input Set : N:\CrF3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

192	acaaggacgc gcccgtact aacaacaagg ccatagccca ggccgagagtc ggagacacgg	9240
193	cattcgctgtt gggaaactac ggcaggccca tcgaggccga cccyccgtac gcgcggccgc	9300
194	gttcaaacag ggcgtgtg ctgcacaggc tggcgagca tgaggaggcg ctgcggacc	9360
195	tgcacaggc agccgtgtg gacccgcga agccgaaacc gaggttctac aaggggatag	9420
196	tgtccggcgtt gatggccagg cacaaggag cgctggcctg cttaaaggc gtgtcaaga	9480
197	ggcacccgg caacccgcac tcacagtcc acgttggat agacttacc gagcttgc	9540
198	ggcacccgcg ggcctcggg gagottgtat cactgcccgc ggacacccgc gagaacccca	9600
199	atgttttgta tgccaggcgg cgcacccctc cggggcttgc caggaggac gaatccatag	9660
200	cgcacccgtca aaaggccggc aaaaaggatc ccaagacgt aaaaaatttg gccccgcgag	9720
201	aaaaggccctt tgacggataa cggacgtatc cgggttccaa aagatagccg cttagaggat	9780
202	cttttttctt ccggcgttca tccgcattatc cggccgttcc cccaaatcg	9840
203	cgttcatatc atctggatcat agacccatcc caccggctt cggccaaact cctcccgag	9900
204	gtccgcgttgc cggccggcgg gggccggccg cagcttctt ttatgtttt agagccgttc	9960
205	tttgcgttcc acctggggc tccgcattatc ctgcgcga tcgatgtatcc tccgggtta	10020
206	cgcgttccccc ggcggccggc cctgtttgtt gggaaatgc gcgtccgcgc cgtccggtc	10080
207	aaggccatcc atcccttctt atccctccaaatc gatcatgtt tctatgtcga ggtccgacc	10140
208	gtttttccgcg acggccgttgc agagcgttgc ttgtatgcgc gatctggacc ttatggaaa	10200
209	gacccgttgc cttagccatc cctcgatcac ttgttgcgc accttgcgcg ggtgaacccgc	10260
210	ctttctgaaa atatccggc agatccgttgc ggagacccgc atcagccctc ctggggaccc	10320
211	cttttacagaa tgcacatggc cgttcttcc cccggggcc cttcaataggc ccctaaaggc	10380
212	accgggttcc ttgttgcata tcatggccgc agcgcacttc tcagtcatgg cgttccgtac	10440
213	gaccggccgtt ctgttcttc cagtatccatc ctgcgcacc cccgatataagg catactatag	10500
214	aacccaaaggc aagttataatc tagccgttgc ttgttgcgc cttatgtatcc ttgggggc	10560
215	ggccggataatc tccaaaggat ttccaaaggc gggccggacca ggcgtatagc aatycgttc	10620
216	acggggcgttgc caaggggcgc gacagatatac tagatgtacgc agtcgatc gcaaatatca	10680
217	ccgttggccgc ggcggaaaaaa agaagcgtatc tgctgtccaa gcaggccgcg cggggagac	10740
218	aggcgcttccaa tccaaagggg gccaaaaaaaaaa tgaaaaaggcatggggc gcaaaaaaaa	10800
219	tggcggccggc caaggccgcac ggctgtatcc cttccggaaaaa gctccggcgcg ctgaaaaagg	10860
220	cggggatcat aacggggaaag gagtttgcgc ccaaaaaaaaaaa gaagggttcc tcggggatcc	10920
221	gacttgcgttgc cgttgcatac taccggggac ggtccgtatcc aggggttcc aaaaaaaaaa	10980
222	tttatagccatc cggccgttcc ttgtcgcattt cggccggcgc tcggccatcc ttggatcac	11040
223	atatttcacgcg ttacatccgc cggccggccaa aaggacccccc tgcatacgca cggggccgg	11100
224	cayctgtccatc tagtgcactt cggccggccgc agatgtccaa tattttaaaaa gaccggccgc	11160
225	ggccggccggg aattttgcata aaaaagagatac gagacgtatcc cttccaaaggc gggccatcc	11220
226	cgtatccatc ctgcggccgtt gtcacgttgc acyccgcac acyccgcacc accctgtccc	11280
227	atatacgccatc aataccatccgc tggccatccgc gaaaaggccgc gttatccatc ttgtatgaat	11340
228	ccgactttgc caccgggttgc accggcgttgc ttgtttttttt attatttgcg cttcttcgt	11400
229	atccgcacgc ttacatccgcgtt ggtccatccgc atcccttgc gttatccatc ctttttcgt	11460
230	ttttccgcgtt actttgtatc gctttgttgc ttctttccgc ttgtatccgc gggccatcc	11520
231	atgttccgcg tgcacggggaa gggccaaatcc ggcggaaatcc tcgagatccatc atccatgcgc	11580
232	ccttattccgcg cttccgtatcc ttcttcgttgc cttccgtatcc cttccgtatcc cccggggat	11640
233	tccttcgttgc aatctccgttgc ttccatgtatcc tttatgtatcc cttccgtatcc cccggggat	11700
234	tccttcgttgc atatctccgttgc ttccatgtatcc tttatgtatcc cttccgtatcc cccggggat	11760
235	tcacccgttgc ttatcatttt gggccgttgc ttgtatccatc cttccgtatcc cccggggat	11820
236	ttgttgcatac agtacggccgtt gcttttgcgc gggccgttgc ttgtatccatc cttccgtatcc	11880
237	cctacgttcc tcgacggatc ccatttcctcc ttgtatccatc tcgaggatcc gtcgtccgt	11940
238	tatccgttgc aggccatccatc cccggggatc ggtccatccgc ttgtatccatc tggatccgc	12000
239	ggaaacccgtt aatataatccatc tccatccgcgc cttttttttt gggggccccc ctttttcgt	12060
240	ggccggccggc aggccgttgc ggtatccatc gttatccatc tttttttttt taaaacttggatccatc	12120

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002

TIME: 17:49:51

Input Set : N:\CrF3\RULE60\10029120.raw  
Output Set: N:\CRF3\01172002\J029120.raw

L:9 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD  
L:7288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120